THE EVOLUTION OF SELF-REGULATED TRANSPOSITION OF TRANSPOSABLE ELEMENTS

B. CHARLESWORTH* AND C. H. LANGLEY†

Department of Biology, The University of Chicago, 1103 East 57th Street, Chicago, Illinois 60637, and †Laboratory of Genetics, National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina 27709

Manuscript received April 26, 1985
Revised copy accepted September 25, 1985

ABSTRACT

This paper examines the conditions under which self-regulated rates of transposition can evolve in populations of transposable elements infecting sexually reproducing hosts. Models of the evolution of both *cis*-acting regulation (transposition immunity) and *trans*-acting regulation (transposition repression) are analyzed. The potential selective advantage to regulation is assumed to be derived from the deleterious effects of mutations associated with the insertion of newly replicated elements. It is shown that both types of regulation can easily evolve in hosts with low rates of genetic recombination per generation, such as bacteria or bacterial plasmids. Conditions are much more restrictive in organisms with relatively free recombination. In haploids, the main selective force promoting regulation is the induction of lethal or sterile mutations by transposition; in diploids, a sufficiently high frequency of dominant lethal or sterile mutations associated with transpositions is required. Data from Drosophila and maize suggest that this requirement can sometimes be met. Coupling of regulatory effects across different families of elements would also aid the evolution of regulation. The selective advantages of restricting transposition to the germ line and of excising elements from somatic cells are discussed.

THERE is a growing theoretical literature on the evolutionary biology and population genetics of transposable elements (OHTA 1981, 1983, 1984, 1985; OHTA and KIMURA 1981; HICKEY 1982; BROOKFIELD 1982, 1985a; CHARLESWORTH and CHARLESWORTH 1983; LANGLEY, BROOKFIELD and KAPLAN 1983; KAPLAN and BROOKFIELD 1983; GINZBURG, BINGHAM and YOO 1984; CHARLESWORTH 1985; KAPLAN, DARDEN and LANGLEY 1985; SLATKIN 1985), stimulated by the increased awareness that repetitive DNA may often be parasitic. Much of this literature is concerned with explaining the main features of the distribution and abundance of transposable elements within host populations, as functions of variables such as transposition and excision rates, selection on host individuals in relation to number of copies of elements, and so on. It is clear that, in sexual organisms, a family of elements can be maintained in a state of stable equilibrium in a large population as a result of a balance between the spread of elements by replicative transposition and the

elimination by selection of individuals carrying such elements. A similar balance can arise if transposition is self-regulated; *i.e.*, if the probability of transposition per element decreases with the number of elements of the same family in the host genome (CHARLESWORTH and CHARLESWORTH 1983; LANGLEY, BROOKFIELD and KAPLAN 1983).

Evidence from Drosophila populations suggests the operation of forces maintaining such an equilibrium, since elements seem to be very rare at individual chromosomal sites which they are capable of occupying (MONTGOMERY and LANGLEY 1983; LEIGH BROWN 1983; C. F. AQUADRO *et al.*, unpublished results). There is abundant evidence for the ability of insertions of transposable elements to cause harmful mutations (SHAPIRO 1983), so that the existence of some selection pressure on host individuals as a function of the number of elements they contain is highly plausible. There is also firm evidence for self-regulation of transposition in prokaryotes (KLECKNER 1981; SHERRATT *et al.* 1983), maize (ROBERTSON 1983; SCHWARTZ 1984) and Drosophila (BREGLIANO and KIDWELL 1983; ENGELS 1983, 1985). Thus, despite a widespread antagonism to the parasitic DNA concept (SYVANEN 1984), there is much evidence in its favor.

The processes of selection on copy number and self-regulation of transposition are not, of course, mutually exclusive. In this paper, we are concerned with explaining regulation of transposition rates in evolutionary terms; *i.e.*, we wish to study processes by which a population of elements which lack such a property come to acquire it through natural selection. We assume that regulation of transposition is an evolutionary response to the harmful effects on the host of the presence of elements in its genome. Our analyses therefore assume the existence of an initial population held at equilibrium between selection and transposition; then, we determine the conditions for the spread of a mutant element that alters the process of transposition in individuals that contain it.

Some work has already been done on the evolution of regulation (DOOLITTLE, KIRKWOOD and DEMPSTER 1984). This work assumes, however, that the host reproduces asexually. It is not clear how a stable equilibrium of copy number can be maintained by the above processes in asexual populations (*c.f.* HICKEY 1982), and it is, in any case, important to investigate the properties of sexual organisms, for which population data are easier to interpret. Accordingly, our models assume sexual reproduction. Mostly, we shall assume diploidy, but extensions to sexual haploids will be discussed. By setting recombination frequencies to near zero, these models should be applicable to such organisms as bacteria and their plasmids, in which genetic exchange between individuals occurs but is probably very rare in nature (HARTL and DYKHUIZEN 1984).

We shall consider two main mechanisms of transposition regulation. The first is transposition immunity (ROBINSON, BENNETT and RICHMOND 1977; SHERRATT *et al.* 1983), in which an element located in a piece of host DNA prevents the insertion of other elements of the same family from inserting in *cis* nearby. There is a considerable amount of evidence for this process in

prokaryote systems (SHERRATT *et al.* 1983); it is unclear whether it occurs in eukaryotes. The second mechanism is what we shall call transposition repression, in which elements produce *trans*-acting repressors of transposition. A good deal is known about this process in prokaryotes (SHERRATT *et al.* 1983; SIMONS and KLECKNER 1983; JOHNSON and REZNIKOFF 1984), and there is evidence for similar processes in eukaryotes (ROBERTSON 1983; SCHWARTZ 1984; BREGLIANO and KIDWELL 1983; ENGELS 1983, 1985). We shall also consider briefly the dynamics of selection on "rogue" elements that ignore signals repressing their transposition, and on mutations of the host genome which affect transposition rates. Our calculations have some formal similarities with those of LEIGH (1970, 1973) on the evolution of mutation rates.

EVOLUTIONARY MODELS: GENERAL PROPERTIES

We assume, throughout, that we are dealing with a discrete generation, sexually reproducing, randomly mating host population of infinite size. The host individuals are assumed to contain a family of transposable elements, for which m chromosomal sites per haploid genome are potentially available for occupation. New elements, generated by replicative transposition, insert at random into unoccupied sites. For algebraic convenience, we shall assume that $2m$ is a very large number, compared with the number of sites normally occupied in a single host individual. The available population data for Drosophila suggest that this is a realistic assumption (MONTGOMERY and LANGLEY 1983; LEIGH BROWN 1983; AQUADRO *et al.*, unpublished results).

We follow the notation of CHARLESWORTH and CHARLESWORTH (1983) and write the probability of replicative transposition of a given element in a single generation of the host as u, and the probability of its loss from the host genome by excision as v. (v may be non-zero.) In the absence of self-regulated transposition, we assume that a stable equilibrium of copy number is maintained in the population by a balance between transpositional increase in copy number and selection against host individuals with high copy numbers. This requires that the fitness (w_n) of individuals as a function of copy number (n) be a concave, decreasing function of n; *i.e.*, that the adverse effect of an additional copy on fitness increases with the number of copies already present (CHARLESWORTH and CHARLESWORTH 1983; CHARLESWORTH 1985; BROOKFIELD 1985b). A biological basis for this assumption is provided by the observation that the viability of chromosomes of *D. melanogaster* which have accumulated harmful mutations with minor effects obeys this rule (MUKAI 1969). If the insertion of a transposable element induces mutations with such minor effects when heterozygous, and if element frequencies are sufficiently low that the mutations are rarely expressed as homozygotes (*c.f.* MONTGOMERY and LANGLEY 1983; LEIGH BROWN 1983; AQUADRO *et al.*, unpublished results), then such a relation between fitness and copy number would be expected (CHARLESWORTH 1985).

If $\bar{n}$ is the mean copy number per individual and $\bar{w}$ is the population mean fitness, then a population in equilibrium under the above assumptions satisfies the balance equation (CHARLESWORTH 1985)

$$\frac{\partial ln\ \bar{w}}{\partial \bar{n}} + u - v = 0. \tag{1}$$

We shall assume that the population is initially genetically homogeneous with respect to the properties of the transposition process and is at equilibrium under (1). Provided $\bar{n}$ is substantially greater than zero and $2m \gg \bar{n}$, the copy number distribution in the population will be approximately normal with variance $\bar{n}$ (CHARLESWORTH and CHARLESWORTH 1983). A mutant element with altered transpositional properties is introduced at low frequency; we shall study the process of spread while it is so rare that only heterozygotes need be considered. We shall also assume that recombination in the host genome is sufficiently frequent that linkage disequilibrium between copies of the mutant element that have transposed to different chromosomal sites can be neglected. [The results of CHARLESWORTH and CHARLESWORTH (1983) justify the assumption of approximate linkage equilibrium among transposable elements under selection.] While the mutant element is still rare, it is then safe to assume that, at most, only a single mutant copy is present per host individual, and this assumption will be made throughout.

Given these assumptions, it is evident that individuals carrying mutant elements will come to be associated with a different mean number of copies of nonmutant elements than the individuals without mutant elements, thus creating a difference in net fitness between mutant and nonmutant individuals that leads to a selection pressure on the mutant element. As shown in the appendices, the process of change in the frequencies of nonmutant elements at each chromosomal site, induced by the presence of the mutant element, can be represented by nonhomogeneous linear equations; it follows that the mean copy number among mutant individuals tends to a constant asymptotic value. Our strategy, therefore, will be to determine this quantity for each class of model and from this to obtain an expression for the selection coefficient for or against the mutation. The details of the derivations are given in the appendices; APPENDIX 1 contains results relevant to all of the models, and the individual cases are treated in subsequent ones.

TRANSPOSITION IMMUNITY

In order to model the evolution of transposition immunity, we assume that a mutant element arises that prevents the insertion of other elements in *cis* into a region ρ map units long, centered on the site of insertion of the mutant. It is not, at present, clear whether or not transposition immunity excludes the insertion of replicates of the element responsible into the region in question; the data of BISHOP and SHERRATT (1984) on Tn*1* suggest that it does not. We shall allow for either possibility in what follows.

Let the total map length of the host genome be L. Assuming that insertions are distributed randomly over the map, the upper limit to any reduction in the rate of transposition of elements present in mutant individuals, due to their inability to insert into the protected region, is $u\rho/2L$, noting that only one-half of possible insertions can occur in *cis* with the mutant element and, hence, be

potentially subject to inhibition. We can write the change in the average rate of transposition of elements present in a mutant individual (including the mutant element itself) as

$$\rho/2L \leq \delta u_a \leq 0. \tag{2}$$

The change in rate of transposition of the mutant element itself, δu_m, satisfies

$$\delta u_a \leq \delta u_m \leq 0. \tag{3}$$

It will be seen from (A.12) of APPENDIX 2 that the alteration in mean copy number associated with the mutant element depends on the quantities

$$\Sigma_1 = \sum_{r_i=0}^{\frac{1}{2}\rho} \frac{1}{(r_i + u)}, \tag{4a}$$

$$\Sigma_2 = \sum_{r_i>\frac{1}{2}\rho}^{\frac{1}{2}} \frac{1}{(r_i + u)}, \tag{4b}$$

where r_i is the frequency of recombination between the site of a mutant element and the ith occupable chromosomal site.

Approximate values for these quantities may be found as follows, assuming that elements are inserted randomly over the entire genome. Let ψ_1 and ψ_2 be the proportions of pairs of loci in the regions $0 < r < \frac{1}{2}\rho$, and $\frac{1}{2}\rho < r < \frac{1}{2}$, respectively, where r represents recombination frequency. Given the assumption of random insertion, we can write

$$\psi_1 = \rho/L, \qquad \psi_2 = 1 - \rho/L. \tag{5}$$

If H_1 and H_2 are the harmonic means of $r + u$ for pairs of loci in these two regions of the genome, we have

$$\frac{1}{H_1} = \frac{1}{m\psi_2}\,\Sigma_1, \qquad \frac{1}{H_2} = \frac{1}{m\psi_2}\,\Sigma_2. \tag{6}$$

Writing $\psi_1(r)$ for the probability density of recombination fraction r for a pair of randomly chosen loci, such that $0 \leq r \leq \frac{1}{2}\rho$, we obtain

$$\frac{1}{H_1} = \int_0^{\frac{1}{2}\rho} \frac{\psi_1(r)}{(r + u)}\,dr. \tag{7}$$

If P_L is the probability that a pair of randomly chosen loci are on the same chromosome, and $\psi_2(r)$ is the probability density of recombination fraction r for such a pair, subject to the restriction $\frac{1}{2}\rho \leq r \leq \frac{1}{2}$, we similarly obtain the relation

$$\frac{1}{H_2} = 2(1 - P_L) + P_L \int_{\frac{1}{2}\rho}^{\frac{1}{2}} \frac{\psi_2(r)}{(r + u)}\,dr. \tag{8}$$

MORTON (1955) has shown that, for a chromosome of sufficient length that obeys KOSAMBI's (1943) mapping function, the probability density of the recombination fraction for a pair of genes drawn from a set of uniformly spaced loci on the same chromosome follows an approximately uniform distribution.

Applying this result to (6), (7) and (8), with the additional assumptions of equal crossover frequencies in the two sexes and l chromosomes of equal length, so that $P_L = 1/l$ and $1 - l_L = 1 - 1/l$, we obtain (after some straightforward calculations) the relations

$$\frac{1}{m}\Sigma_1 \simeq 2\, ln\left(1 + \frac{\rho}{2u}\right)\Big/ L \tag{9}$$

$$\frac{1}{m}\Sigma_2 \simeq 2\left\{1 - \frac{(1 + ln\ \rho)}{l}\right\} \qquad (\text{if } \rho \ll 1). \tag{10}$$

Substituting into (A.12), we obtain the following expression for the reduction in the mean copy number of mutant individuals compared with that for nonmutants:

$$-\delta\bar{n} \simeq \bar{n}\left\{\frac{u}{L}\, ln\left(1 + \frac{\rho}{2u}\right) - \delta u_a\left(1 - \frac{1}{l}\,[1 + ln\ \rho]\right)\right\}. \tag{11}$$

The first term on the right-hand side represents the reduction in mean copy number due to the fact that no elements can insert in *cis* with the mutant element; the second term corresponds to the reduction associated with the possible overall lower rate of transposition in mutant individuals. Since $-\delta u_a < u\rho/2L$, from (2), and since ρ is likely to be $\ll 1$, the first term usually must dominate, and the second term can be neglected.

Substituting into (A.16), we obtain the selection coefficient on the mutant element as

$$s \simeq \frac{\bar{n}u(u - v)}{L}\, ln\left(1 + \frac{\rho}{2u}\right) + \delta u_m. \tag{12}$$

If there is no reduction in the transposition rate of the mutant element associated with its ability to inhibit insertions in its vicinity, then $\delta u_m = 0$, and such an element will always be favored. It is clear from (12), however, that its selective advantage is inversely proportional to the total genomic map length and is of the order of $\bar{n}u^2$, unless ρ/u is exceedingly large. If $\bar{n} \simeq 50$ and u is of order 10^{-4}, as seems plausible for many Drosophila elements (CHARLESWORTH and CHARLESWORTH 1983), then s must be of order $10^{-7}/L$ and can, therefore, take a substantial value only if the genomic map length is very small (see the later treatment of this case).

If the mutant elements have their transposition rate reduced by the maximal amount $u\rho/2L$ [(2) and (3)], they will experience a selective advantage only if

$$\bar{n}(u - v)ln\left(1 + \frac{\rho}{2u}\right) > \tfrac{1}{2}\rho. \tag{13}$$

This places a severe restriction on the size of the region ρ over which transposition immunity can extend. For example, if $\bar{n} = 50$, $u = 10^{-4}$ and $v = 5 \times 10^{-5}$, then ρ would have to be approximately 0.01 for this condition to be satisfied; *i.e.*, there would be selection only for mutants that restrict insertion in a region about 1 map unit long. If transposition rates were changed by a

given factor, the length of the protected region would change roughly proportionately.

TRANSPOSITION REPRESSION

In order to model the evolution of transposition repression, we proceed much as before, but assume that the mutant element affects the rate of transposition of all elements in the same individual (including itself), regardless of their chromosomal location. We shall consider three versions of this model: a simple version that uses only the above assumptions, a version in which repression is conditional on the number of elements in the same individual as the mutant, and a version in which transposition induces dominant lethal or sterile mutations with a certain frequency.

A simple model: Here, we assume that the mutant element changes the rate of transposition of all elements in a mutant individual by δu_a; we now have $\delta u_m = \delta u_a$. The consequences of this assumption are developed in APPENDIX 3. Equations (A.18) and (A.19) provide expressions for $\delta\bar{n}$ and s in this case. $\delta\bar{n}$ now depends on the harmonic mean, H, of $r + u$, where r is taken over all randomly chosen pairs of loci. It follows from (A.19) that transposition repression will be favored only if

$$\frac{\bar{n}(u - v)}{2H} > 1. \tag{14}$$

Using the same methods as for transposition immunity, H for the case of equal crossover frequencies in the two sexes is given by

$$\frac{1}{H} = 2(1 - P_L) + P_L \int_0^{1/2} \frac{\psi(r)}{(r + u)}\, dr. \tag{15}$$

Assuming l chromosomes of equal length and a uniform distribution of recombination frequencies between random pairs of loci, we obtain

$$\frac{1}{2H} = \left(1 - \frac{1}{l}\right) + \frac{1}{l}\, ln\left(1 + \frac{1}{2u}\right). \tag{16}$$

With completely free recombination ($l = \infty$), $\frac{1}{2H} = 1$. With finite l, H is weakly dependent on u, due to the logarithmic term in (16). It is more strongly dependent on l, as may be seen in Table 1. Nevertheless, the multiplier of $\bar{n}(u - v)$ in (14) approaches only a value of ten even for an organism with a single pair of chromosomes, so that $\bar{n}(u - v) > 0.1$ is a necessary condition for the evolution of transposition repression with these assumptions.

The effect of sex differences in crossing over on these conditions is of interest, in view of the absence of crossing over in males of many species of Diptera, including Drosophila. In the case of *D. melanogaster*, one-fifth of the genome is on the X chromosome and most of the remainder is approximately equally divided between the two major autosomes. P_L is, thus, approximately $(1/5)^2 + 2(2/5)^2 = 0.36$. On account of the lack of crossing over in males, r

TABLE 1

Dependence of $1/(2H)$ on the number of chromosomes (l) and the transposition rate (u)

	l		
u	1	10	100
10^{-7}	15.42	2.44	1.14
10^{-6}	13.12	2.21	1.12
10^{-5}	10.82	1.98	1.10
10^{-4}	8.52	1.75	1.08

in (15) is replaced with $\frac{1}{2}r$, and we obtain

$$\frac{1}{2H} \simeq 0.64 + 0.72\ ln\left(1 + \frac{1}{4u}\right). \tag{17}$$

For $u = 10^{-5}$, $\frac{1}{2H}$ is now 7.93, compared with 4.54 with male crossing over. This difference does not materially affect the chance of condition (14) being satisfied.

These calculations do not take account of the well-known inequalities in crossover frequencies per unit of DNA in Drosophila (LINDSLEY and SANDLER 1977). Since transposable elements will presumably be inserted more nearly at random into the physical rather than the genetic map, our assumption of a uniform distribution with respect to genetic maps will tend to overestimate H. This question will be further explored in the DISCUSSION.

Conditional repression of transposition: It is intuitively appealing to suppose that a mutant element that represses transposition only when present in an individual whose copy number is large compared with the population mean would experience a greater advantage than in the previous model, since it would suffer a loss only by reducing its own transposition in circumstances when the gain from repressing others' is high. The properties of the transmission of cytotype in the $P - M$ system of hybrid dysgenesis suggest that a threshold number of elements per genome may be required to suppress transposition of P elements (ENGELS 1983, 1985).

The above model can be extended as follows to incorporate this effect. We assume that there is a threshold number of elements per individual, n_c, such that transposition is not repressed when $n \leq n_c$, but is completely repressed for $n > n_c$. Using the methods described in APPENDIX 4, we find from (A.24) that a mutation causing this behavior will spread if

$$\frac{\bar{n}(u - v)F}{2H} > 1, \tag{18}$$

where the factor F measures the advantage over the simple model:

$$F = \frac{(1 - \Phi_c + \bar{n}^{1/2}\phi_c)}{(1 - \Phi_c)}.$$

The quantities ϕ_c and Φ_c are the values of the ordinate and cumulative probability of the normal distribution at n_c, assuming copy number to follow a normal distribution with mean and variance $\bar{n}$. For $n_c \gg \bar{n}$, the asymptotic expansion of the normal probability integral gives

$$1 - \Phi_c \simeq \phi_c \bar{n}^{1/2}(n_c - \bar{n})^{-1}$$

so that

$$F \simeq 1 + n_c - \bar{n}. \tag{19}$$

(This approximation has an error of <12% for values of n_c such that $(n_c - \bar{n})\bar{n}^{-1/2} > 2$ and always provides an underestimate.)

When the population mean is 50, the multiplier F is approximately 15 if the mutant element represses transposition in individuals with a copy number of over 64, who comprise 2.5% of the population. Although it is always possible to achieve a selective advantage for repression with this model, simply by taking n_c sufficiently large, it is evident from (A.24) that the selective coefficient s is proportional to $u(1 - \Phi_c)$, which approaches zero rapidly as n_c increases. In the above example, $1 - \Phi_c \simeq 0.025$ for $n_c = 64$ and 0.0014 for $n_c = 71$. It is evident that, for reasonable values of $\bar{n}$, $u - v$ and H, any positive selection pressure for mutant elements causing transposition repression will be very small.

Dominant lethals and transposition repression: The above models have assumed that all mutations induced by transposition have similar, minor effects on fitness when heterozygous. The evidence from studies of viability mutations in Drosophila suggests that this, indeed, is true of the vast bulk of mutations, including recessive lethals (SIMMONS and CROW 1977). It is possible, however, that some transpositional events have much more drastic effects on fitness; *e.g.*, because of the induction of chromosome rearrangements which segregate abnormally. Prokaryote transposons are well known to induce rearrangements during transposition (KLECKNER 1981), and the *Ac* − *Ds* system of maize induces chromosome breaks (MCCLINTOCK 1956). In *D. melanogaster*, the mobilization of the *P* element in dysgenic crosses is associated with chromosome rearrangements (ENGELS and PRESTON 1984) and chromosome loss (W. BENZ and W. R. ENGELS, personal communication). The *L* factor of LIM (1979), which induces chromosome rearrangements at high frequency (LIM *et al.* 1983), is apparently a highly mobile *gypsy* element (B. H. JUDD and J. JACK, personal communication).

In order to model this possibility, we shall assume that a fraction π of insertions of newly transposed elements into a chromosome are associated with lethality or sterility of the progeny. The expected number of such mutations in the progeny of an individual with copy number n is $n\pi u$. If the mean copy number in the initial population, before introduction of a mutant element, is $\bar{n}$, then the reduction in mean copy number in the next generation due to elimination of such dominant mutations is $\bar{n}\pi u$, and the increase due to non-

lethal transpositions is $\bar{n}(1 - \pi)u$. The net change in mean due to transposition and selection is thus $\bar{n}(1 - 2\pi)u - \bar{n}v$. Taking into account the selective effects of small mutations in the same way as before, the balance equation (1) becomes

$$\frac{\partial ln\ \bar{w}}{\partial \bar{n}} + (1 - 2\pi)u - v = 0. \tag{20}$$

The transposition rate u is thus replaced by an effective rate $\tilde{u} = (1 - 2\pi)u$; this, and the change in it caused by the mutation, $\delta\tilde{u}_a = (1 - 2\pi)\delta u_a$, replace the previous quantities throughout the calculations (APPENDIX 5).

It follows from (A.25) that a mutant element causing transposition repression will be favored if

$$\frac{\bar{n}(\tilde{u} - v)}{2\tilde{H}} + \frac{\bar{n}\pi}{2(1 - 2\pi)} > 1, \tag{21}$$

where $\tilde{H}$ is the harmonic mean of $r + \tilde{u}$ for random pairs of loci.

This differs materially from the condition for the simple model [(14)] in the term $\bar{n}\pi/2(1 - 2\pi)$, which arises from the immediate advantage to the mutation of reducing the frequency of lethal or sterile progeny. The condition for an advantage to transposition repression from this effect alone is

$$\bar{n}\pi > 2(1 - 2\pi) \quad \text{or} \quad \pi > 2/(\bar{n} + 4). \tag{22}$$

With a mean copy number of 50, π would have to exceed 3.7% for this effect alone to confer an advantage to repression.

Dominant lethal mutations could, of course, play a role in the evolution of transposition immunity. As in (2), the upper limit to the reduction in transposition rate $-\ \delta u_a$ is $u\rho/2L$, and the selection coefficient on a mutant element is $<u\rho\bar{n}/\{4(1 - 2\pi)L\}$. Since transposition immunity can probably extend only across short physical distances (SHERRATT *et al.* 1983), the ratio ρ/L is likely to be very small, except in organisms with small genomes.

ADDITIONAL FACTORS

In this section, we shall consider some additional factors that may affect selection for regulation of transposition.

Haploidy: The above models apply to diploid organisms. The treatment of sexually reproducing species with an extended, single-celled, haploid vegetative phase (followed by the fusion of gametes to produce a transient zygote that undergoes meiosis to restart the vegetative phase) is very similar and applies to organisms of genetic importance such as yeast and Chlamydomonas. The details of the necessary modifications are given in APPENDIX 6. Transposition is assumed to occur in the vegetative phase.

For the transposition immunity model, it follows from (A.28) that (12) of the text is changed only by replacing mean copy number $\bar{n}$ for diploid individuals by the mean for haploids and by multiplying the first term on the

right-hand side by two:

$$s \simeq \frac{2\bar{n}u(u - v)}{L} ln\left(1 + \frac{\rho}{2u}\right) + \delta u_m. \tag{23}$$

For equivalent values of $\bar{n}$, conditions are thus somewhat lighter than for diploids if $\delta u_m = 0$. If, however, $-\delta u_m$ takes its maximum value, this is now $\rho u/L$ instead of $\rho uL/2L$, since all insertions must occur in *cis* with the mutation. Thus, the condition for an advantage to the mutation [(13)] is unchanged.

With transposition repression, (A.29) and (A.30) imply that the quantity $1/H$ in (15) is replaced by

$$\frac{1}{H^*} \simeq 2(1 - P_L) + P_L \int_0^{1/2} \frac{(1 - r)\psi(r)}{(r + u)}\, dr. \tag{24}$$

For the case of equal crossing over in the two sexes and a uniform distribution of r for genes on the same chromosome, this becomes

$$\frac{1}{H^*} \simeq 2(1 - P_L) + 2P_L\left\{ln\left(1 + \frac{1}{2u}\right) - 1/2\right\}. \tag{25}$$

Thus, $1/H^*$ is slightly smaller than $1/H$ in (15) and (16). This effect is trivial, however, in comparison with the fact that $2H$ in condition (14) is replaced by H^* rather than $2H^*$.

The dominant lethality model is replaced by one in which π now represents the probability that an insertion of a newly replicated element induces a lethal or sterile mutation in a haploid cell. We might well expect π here to take a much larger value than with diploidy, since mutation events that would be classed as recessive in diploids are now included. Furthermore, if their effects are expressed before gamete formation, all the associated genetic deaths cause the elimination of the mutant element, instead of one-half as in the case of a diploid, multicellular organism, where only one-half of the lethal insertions will be transmitted in *cis* with the mutation. The direct selective advantage to the mutation is now $-\bar{n}\delta u_a$, and condition (22) is modified to

$$\bar{n}\pi > (1 - 2\pi) \quad \text{or} \quad \pi > 1/(\bar{n} + 2) \tag{26}$$

Restricted recombination: It seems clear from population genetic data that genetic exchange is very infrequent in natural populations of bacteria (HARTL and DYKHUIZEN 1984). Similarly, estimates of the rate of transfer of plasmids between bacterial cells are of the order of 10^{-8} or less per cell per minute (LEVIN, STEWART and RICE 1979; FRETER, FRETER and BRICKNER 1983). At least to a first approximation, the above results on haploids can be extended to such organisms by taking recombination frequencies to the limit of zero in the equations for change in element frequencies. For transposition immunity, we now find that Σ_1 of (A.28) becomes

$$\Sigma_1 = \lim_{r_i \to 0} \sum_{i=0}^{1/2\rho} \frac{1}{(r_i + u)} = \frac{m\rho}{uL},$$

so that

$$\delta\bar{n} \simeq -\bar{n}\rho/L \tag{27}$$

$$s \simeq \frac{\bar{n}\rho(u - v)}{L} + \delta u_m. \tag{28}$$

The selection coefficient of the mutation is always much larger than that given by (23) for relatively free recombination. If $-\delta u_m$ takes its maximum value of $u\rho/L$, the mutation will experience an advantage if

$$\bar{n}(u - v) > u. \tag{29}$$

This condition is always satisfied if the mean copy number per cell exceeds one, which is a very light condition.

A similar calculation can be carried out for the simple model of transposition repression. We obtain

$$\delta\bar{n} \simeq \frac{\bar{n}\delta u_a}{u} \tag{30}$$

$$s \simeq -\delta u_a\left\{\frac{\bar{n}(u - v)}{u} - 1\right\}. \tag{31}$$

Again, the selection coefficient of the mutation is much greater than previously, and the condition for its spread is identical with (29).

Selection on sensitivity to repression: Since transposable elements maintain themselves in the host population by virtue of their ability to replicate within the host genome, it is intuitively obvious that a mutant element that ignores repressor molecules will obtain a selective advantage. That this is indeed so can be seen by the following argument, which will be presented for the case of a diploid with a frequency π of dominant lethal or sterile insertions. Let y_i be the frequency at the ith occupable site of a mutant element that modifies its own transposition rate by δu_m but leaves the transposition rate of other elements in the same genome unchanged. For small y_i, using the standard formulation for change in element frequencies, we obtain

$$y_i' = y_i\left\{\frac{\partial ln\ \bar{w}}{\partial\bar{n}} + (1 - 2\pi)(u + \delta u_m) - v\right\}. \tag{32}$$

If the initial population is at equilibrium, the balance equation (20) will be satisfied, and $\Sigma y_i' > \Sigma y_i$ if

$$(1 - 2\pi)\delta u_m > 0. \tag{33}$$

Since the balance equation cannot be satisfied if $1 - 2\pi < 0$, this implies that a mutant element with an increased rate of transposition will always be favored.

Selection on host repressor genes: The converse is expected for genes of the host: a host mutation that represses transposition will be favored, since the fitness of host individuals will be improved as a consequence. Selection on a host gene can be modeled in exactly the same way as selection on a mutant element in the simple transposition repression model, except that now there is

no cost involved in reducing transposition rate. Allowing for dominant lethality or sterility, the selection coefficient is given by the analog of (A.25),

$$s \simeq -\delta u_a \left\{ \frac{\bar{n}(u - v)}{2\hat{H}} + \frac{\bar{n}\pi}{2(1 - 2\pi)} \right\}. \tag{34}$$

Provided $\delta u_a < 0$, s is always positive, as expected.

DISCUSSION

The strength of selection for transposition regulation in eukaryotes: Our analyses show that most mechanisms of evolution of self-regulated transposition produce, at best, a rather weak selective advantage in species with relatively free recombination. A mutant element conferring transposition immunity will always be favored if it suffers no loss in its own rate of transposition [(12)], but its selective advantage is likely to be of the order of 10^{-7} divided by the total genomic map length, for a typical Drosophila element. The evolution of transposition repression by the simple model is virtually impossible with free recombination [(14)–(17)]. The model of conditional repression can produce a selective advantage if the threshold number of elements required for repression to operate is sufficiently high, but the magnitude of this selective advantage is vanishingly small [(18)].

The most promising model for the evolution of self-regulation in diploid eukaryotes invokes dominant lethal or sterile mutations associated with transposition [(21) and (22)]. If this force is to be effective, it requires that the probability of a lethal or sterile mutation per transposition event be sufficiently high: 3.7% for a mean copy number of 50 [(22)]. It is difficult at present to tell whether or not this requirement can be satisfied. For Drosophila, the most informative data are those of ENGELS and PRESTON (1984), who have studied the induction of *X*-chromosome rearrangements by *P* elements in hybrid dysgenic males. Their results suggest that chromosome breaks are induced at random during the dysgenic mobilization of *P* elements and that the number of breaks per *X* chromosome follows a Poisson distribution with a mean of 0.12. The probability of a single break is, accordingly, 0.11; such a break presumably would be dominant lethal and nonrecoverable under the experimental conditions. Recent results of W. BENZ and W. R. ENGELS (personal communication) indicate that *P*-element-bearing *X* chromosomes are lost from dysgenic males at a frequency of at least 1%.

Data on the number of new insertions of *P* elements in dysgenic crosses give values of 0.8–0.1 insertions per *X* per generation (ENGELS 1983). Thus, the probability of a single break per insertion is between 0.19 and 1 for the *X* chromosome. This is probably an overestimate, because only viable chromosomes can be scored for insertions. [On the other hand, certain classes of rearrangements that lead to genetic death, such as *X*-autosome translocations, were not scored (ENGELS and PRESTON 1984)]. All in all, the data suggest that the probability that *P*-element transposition is associated with dominant lethal or sterile chromosome breaks is sufficiently high that selection for transposition repression would be possible.

Additional information on this point is provided by the results of ENGELS (1979) and KIDWELL (1984), drawn to our attention by W. R. ENGELS. ENGELS (1979, Table 1) showed that *P*-element-bearing chromosomes suffer a considerable reduction in probability of recovery from dysgenic compared with nondysgenic males. The simplest explanation of this fact is that dominant lethal chromosome breaks are induced by transposition-related excisions of *P* elements. KIDWELL (1984) found a lowered frequency of hatching of embryos from dysgenic compared with nondysgenic parents, presumably because of dominant lethal events. The magnitude of both these effects is sufficiently high that the criterion for selection for repression would be met easily if the frequency of transposition on the second and third chromosome arms is comparable with that for the *X* in dysgenic crosses.

An important additional factor that may promote the evolution of repression in the case of the *P* and *I* elements of *D. melanogaster* is the sterility associated with mobilization of elements in dysgenic crosses. Although the cause of the sterility syndromes is far from clear (BREGLIANO and KIDWELL 1983; ENGELS 1983), formally one can regard the death of germ cells associated with the mobilization of *P* elements as providing a selective advantage to regulation similar to that in the haploid model of lethality [(26)], where π is now the probability of death of a germ cell per transposition event. Similarly, the reduced hatchability of eggs with $I - R$ dysgenesis can be treated as analogous to the dominant lethal model discussed above. A relatively simple scenario for the evolution of self-regulation of *P* and *I* elements can thus be imagined, with an ancestral population experiencing high rates of transposition and associated fertility loss and with the elements being held in check according to the balance equation (20). (It is important to note that dominant lethality or sterility cannot, in itself, maintain a stable equilibrium of copy number.) The conditions for invasion by mutant elements repressing transposition rates might then be met relatively easily. It is not clear, however, whether the complex rules governing cytotype transmission (BREGLIANO and KIDWELL 1983; ENGELS 1983) would be compatible with our conditions, and this requires further theoretical investigation.

Further evidence concerning the validity of the dominant lethal model comes from the *Ac/Ds* transposable element system of maize, which is associated with frequent chromosome breakage (FEDOROFF 1983). MCCLINTOCK (1956, pp. 211–212) showed that the transposition of *Ds* is frequently associated with highly deleterious effects on the viability of maize kernels; at least some of these effects are not due directly to deletions of genetic material. This system, therefore, is likely to satisfy our condition and is known to be under repressive control by *Ac* in dividing cells (SCHWARTZ 1984).

It seems probable *a priori* that the insertion of transposable elements will have a much higher chance of inducing a directly expressed lethal or sterile mutation in haploid eukaryotes than in diploids. In the yeast *Saccharomyces cerevisiae*, there are around 30 copies per cell of the element *Ty* (ROEDER and FINK 1983), so that the chance of a lethal insertion need be only 3.1% for regulation to be favored [(26)]. Unfortunately, transpositions of *Ty* are nor-

mally detected by selection of associated mutations, so that it is unknown how frequently insertions induce lethals under normal conditions (ROEDER and FINK 1983).

Even under the most favorable model—the one with direct effects of lethality—the selective advantage of a mutation causing transposition is, at most, of the same order as the change in transposition rate, δu_a [(A.25)]. Contemporary transposition rates are usually of the order of 10^{-4} or less (SHAPIRO 1983); thus, the selection coefficient of a mutation perturbing them is likely to be even smaller. This raises the question of how effective such small selection coefficients are likely to be, to which a definite answer cannot be given. A selection coefficient of less than the reciprocal of the species' effective population size is unlikely to be effective against random genetic drift (KIMURA 1983, pp. 43–46).

Taken together, these considerations appear to rule out the simple model of transposition repression and conditional transposition repression in organisms with fairly free recombination, assuming that different families of transposable elements are under independent regulation. The models could be saved if there were coupling of regulatory effects between different element families, so that a mutation in a member of one family could affect the rate of transposition of several other families. If such a process could occur, then $\bar{n}$ in (14) might change by up to two orders of magnitude, enabling the condition for spread to be satisfied for reasonable values of u and v. Despite some evidence for mobilization of *copia* in *P–M* dysgenesis (RUBIN, KIDWELL and BINGHAM 1982), it is uncertain at present whether such coupling exists.

Genomes with restricted recombination: Our results show that genomes with near-zero recombination are far more favorable to the evolution of both transposition immunity and transposition repression than the organisms with fairly free recombination discussed above: there is a simple requirement in a haploid species for mean copy number per individual to exceed one [condition (29)]. This condition assumes that the frequency of recombination between all pairs of sites per generation is much less than the transposition rate. It should be borne in mind that this result is only approximate, since linkage disequilibrium between elements at different sites has been neglected, but may be significant when genetic exchange is very infrequent. If our results are applicable, there is little difficulty in understanding how self-regulation of transposition can be advantageous in bacteria and bacterial plasmids, with their apparently very low rates of genetic exchange per cell generation (HARTL and DYKHUIZEN 1984).

It might be thought that these results on low rates of transposition could illuminate the situation in organisms such as Drosophila, where a large fraction of the total DNA is contained in the proximal heterochromatin, which contains abundant representatives of various transposable element families (RUBIN 1983). Crossing over is extremely infrequent in centric heterochromatin (SCHALET and LEFEVRE 1976; CARPENTER and BAKER 1982); therefore, it might be imagined that the existence of large stretches of such chromatin with the potentiality for harboring transposable elements would be favorable to the

evolution of regulated transposition. But proximal heterochromatin is largely devoid of functional gene loci (MULLER and PAINTER 1932; HILLIKER, APPELS and SCHALET 1980); therefore, insertions of transposable elements into heterochromatin are unlikely to cause mutations with selective effects, unless dominant lethal or sterile chromosomal breaks are induced. Hence, a mutant element that is located in heterochromatin will not benefit by inhibiting the insertion of other elements into its immediate neighborhood, since most of these insertions will be of little selective effect. Heterochromatin may have an indirect effect through increasing the mean copy number over that which would be maintained by selection/transposition balance in the absence of a reservoir of sites where insertions are neutral (CHARLESWORTH 1985), but this will have only a minor effect on the conditions for transposition regulation.

The reduced rates of recombination in euchromatin adjacent to the proximal heterochromatin (LINDSLEY and SANDLER 1977) will tend to increase the advantage of both transposition immunity and repression over our values, but it is unlikely that this would reduce the harmonic mean recombination frequency by more than two- or three-fold.

Effects of inbreeding: In species with a high level of close inbreeding, such as self-fertilizing plants, recessive lethal or sterile mutations will be exposed to selection almost immediately after their occurrence. Recombination is also effectively very restricted, due to high levels of homozygosity. For these reasons, regulation of transposition should be selected for more effectively in such species. Evidence on this point is currently lacking.

Evolutionarily stable strategy (ESS) transposition rates: So far, we have written as if we were considering only the invasion of a population lacking any form of regulation by mutations causing regulation. But our models of transposition repression can be extended to yield conditions for the invasion of a genetically homogeneous population by a mutation with an effect on transposition rate in either direction. Therefore, we can determine the ESS (MAYNARD SMITH 1982) for transposition rate by asking what transposition rate for the population is stable to invasion by mutations which cause small perturbations. For example, in the diploid dominant mutation model, (22) tells us that the ESS transposition rate is such that $\bar{n} = 2(1 - 2\pi)/\pi$. From (20), we know that $\bar{n}$ at equilibrium is an increasing function of transposition rate u, given that $\partial^2\bar{w}/\partial\bar{n}^2 < 0$, so that the ESS value of $u(u^*)$ satisfies (20) with $\bar{n} = 2(1 - 2\pi)/\pi$. The exact value of u depends on the fitness function assumed. For $w_n = \exp - \frac{1}{2}tn^2$, we find that $u^* = v + 2t(1 - 2\pi)/\pi$.

The utility of such calculations of the outcome of selection is weakened by the existence of possible pleiotropic costs to changes in the process of transposition, and by the existence of a selection pressure on hosts to reduce transposition rates and on elements to ignore repressors of transposition (see the previous section). If an evolutionary race is not to continue indefinitely between selection for and against increased transposition rates, the final equilibrium will be determined by the values of the costs and benefits for all of the parties concerned. As can be seen from (32) and (33), the selection coefficient on a mutant element for which the only effect is to increase its own transpo-

sition rate by δu_m is equal to δu_m. For genomes with relatively free recombination, our results show that the selection coefficients on genes causing transposition immunity or transposition repression (in the absence of dominant lethality or sterility) are probably always far smaller than this. Given comparable pleiotropic costs of changes, we would expect, therefore, selection for "rogue" elements that ignore repressor signals to overcome selection in favor of self-regulation of this kind. Selection on host genes that repress transposition is also weak in these models. With lethality and sufficiently large π, however, a comparable selection pressure for and against transposition repression will be exerted, even with fairly free recombination. With restricted recombination, each model of selection for regulation can exert a similar selection pressure to that exerted on rogue elements, under suitable conditions. We would predict, therefore, that transposition immunity and simple transposition repression will tend to evolve only in organisms with small genomes and restricted recombination, and that transposition repression will evolve in organisms with free recombination only if the frequency of immediately expressed lethal or sterile mutations is sufficiently high.

Somatic repression of transposition and somatic elimination: In organisms with a strict division between germ line and soma, transposition in somatic cells confers no competitive advantage to transposable elements, because there is no possibility of transmission to the next generation; it may be positively disadvantageous because of the induction of somatic mutations that reduce the host's chance of survival to reproductive maturity. Thus, there is a strong selective advantage to elements which repress transposition in somatic tissues. [This point has been made independently by J. F. CROW and by J. A. SVED (personal communications).] A striking feature of Drosophila transposable elements is their lack of somatic transposition; this is true of both the *I* and *P* hybrid dysgenesis elements and of the foldback element (RUBIN 1983; ENGELS 1983; BREGLIANO and KIDWELL 1983), as evidenced by the somatic stability of germinally unstable mutations caused by insertions and by the lack of somatic effects in dysgenesis. The high degree of somatic instability of maize elements (FEDEROFF 1983) is consistent with this, since the soma/germ plasm distinction is blurred in plants.

This argument can be extended to suggest that they may be a selective advantage to the excision of elements from somatic tissues. This could either be a means of reducing the burden of mutational effects associated with element insertions or be a way of accelerating DNA replication. In each case, selection could act either at the level of host or at the level of the elements themselves. The behavior of the element Tc*1* of *Caenorhabditis elegans*, which is specifically excised from somatic cells (EMMONS and YESNER 1984), fits this prediction; Tetrahymena also undergoes elimination of moderately repeated DNA sequences from its "somatic" macronucleus (YOKOYAMA and YAO 1982). It is not clear whether the elimination of supernumerary chromosomes from somatic tissues of gall flies (WHITE 1973, pp. 523–538), the elimination of highly repeated telomeric DNA from Ascaris somatic cells (ROTH and MORITZ 1981) or chromatin elimination in Cyclops (BEERMANN and MEYER 1980) have

a similar explanation, since the tandemly repeated sequences involved may have a nonparasitic role in these organisms.

This work was supported by an SERC Visiting Fellowship to C.H.L. at the University of Sussex, England. We thank W. R. ENGELS, M. R. ROSE, B. S. WEIR and two reviewers for their comments on the manuscript.

LITERATURE CITED

BEERMANN, S. and S. F. MEYER, 1980 Chromatin rings as products of chromatin diminution in *Cyclops*. Chromosoma **77:** 277–283.

BISHOP, R. and D. SHERRATT, 1984 Transposon Tn*1* intra-molecular transposition. Mol. Gen. Genet. **196:** 117–122.

BREGLIANO, J.-C. and M. G. KIDWELL, 1983 Hybrid dysgenesis determinants. pp. 363–410. In: *Mobile Genetic Elements*, Edited by J. A. SHAPIRO. Academic Press, New York.

BROOKFIELD, J. F. Y., 1982 Interspersed repetitive DNA sequences are unlikely to be parasitic. J. Theor. Biol. **94:** 281–299.

BROOKFIELD, J. F. Y., 1985a A model for DNA sequence evolution within transposable element families. Genetics **112:** 393–408.

BROOKFIELD, J. F. Y., 1985b A simple analytical model for the spread of independent transposable elements of equal effect. J. Theor. Biol. In press.

BULMER, M. G., 1980 *The Mathematical Theory of Quantitative Genetics*. Oxford University Press, Oxford.

CARPENTER, A. T. C. and B. S. BAKER, 1982 On the control of the distribution of meiotic exchange in *Drosophila melanogaster*. Genetics **101:** 81–89.

CHARLESWORTH, B., 1985 The population genetics of transposable elements. In: *Proceeding of the Oji International Seminar*, Edited by T. OHTA and K.-I. AOKI. In press.

CHARLESWORTH, B. and D. CHARLESWORTH, 1983 The population dynamics of transposable elements. Genet. Res. **42:** 1–27.

DOOLITTLE, W. F., T. B. L. KIRKWOOD and M. A. H. DEMPSTER, 1984 Selfish DNA with self-restraint. Nature **307:** 501–502.

EMMONS, S. W. and L. YESNER, 1984 High-frequency excision of transposable element Tc*1* in the nematode *Caenorhabditis elegans*. Cell **36:** 599–605.

ENGELS, W. R. 1979 Germ line aberrations associated with a case of hybrid dysgenesis in *Drosophila melanogaster* males. Genet. Res. **33:** 137–146.

ENGELS, W. R., 1983 The *P* family of transposable elements in *Drosophila*. Annu. Rev. Genet. **17:** 315–344.

ENGELS, W. R., 1985 On the evolution and population genetics of hybrid dysgenesis-causing transposable elements in Drosophila. Proc. Roy. Soc. Lond. (Biol.). In press.

ENGELS, W. R. and C. R. PRESTON, 1984 Formation of chromosome rearrangements by *P* factors in Drosophila. Genetics **107:** 657–678.

FEDOROFF, N. V., 1983 Controlling elements in maize. pp. 1–63. In: *Mobile Genetic Elements*, Edited by J. A. SHAPIRO. Academic Press, New York.

FRETER, R., R. R. FRETER and H. BRICKNER, 1983 Experimental and mathematical models of *Escherichia coli* plasmid transfer in vivo and in vitro. Infect. Immun. **39:** 60–84.

GINZBURG, L. R., P. M. BINGHAM and S. YOO, 1984 On the theory of speciation induced by transposable elements. Genetics **107:** 331–341.

HARTL, D. L. and D. E. DYKHUIZEN, 1984 The population genetics of *Escherichia coli*. Annu. Rev. Genet. **18:** 31–68.

HICKEY, D. A., 1982 Selfish DNA: a sexually-transmitted nuclear parasite. Genetics **101:** 519–531.

HILLIKER, A. J., R. APPELS and A. SCHALET, 1980 The genetic analysis of *D. melanogaster* heterochromatin. Cell **21:** 607–619.

JOHNSON, R. C. and W. S. REZNIKOFF, 1984 Copy number control of Tn*5* transposition. Genetics **107:** 9–18.

KAPLAN, N. L. and J. F. Y. BROOKFIELD, 1983 The effect on homozygosity of selective differences between sites of transposable elements. Theor. Pop. Biol. **23:** 273–280.

KAPLAN, N. L., T. DARDEN and C. H. LANGLEY, 1985 Evolution and extinction of transposable elements in Mendelian populations. Genetics **109:** 459–480.

KIDWELL, M. G., 1984 Hybrid dysgenesis in *Drosophila melanogaster*: partial sterility associated with embryo lethality in the *P-M* system. Genet. Res. **44:** 11–28.

KIMURA, M., 1983 *The Neutral Theory of Molecular Evolution*. Cambridge University Press, Cambridge.

KLECKNER, N., 1981 Transposable elements in prokaryotes. Annu. Rev. Genet. **15:** 341–404.

KOSAMBI, D. D., 1943 The estimation of map distances from recombination values. Ann. Eugen. (Lond.) **12:** 172–175.

LANGLEY, C. H., J. F. Y. BROOKFIELD and N. L. KAPLAN, 1983 Transposable elements in Mendelian populations. I. A theory. Genetics **104:** 457–472.

LEIGH, E. G., 1970 Natural selection and mutability. Am. Nat. **104:** 301–305.

LEIGH, E. G., 1973 The evolution of mutation rates. Genetics **73** (Suppl): s1–s18.

LEIGH BROWN, A. J., 1983 Variation at the 87A heat shock locus in *Drosophila melanogaster*. Proc. Natl. Acad. Sci. USA **80:** 5350–5354.

LEVIN, B. R., F. M. STEWART and V. A. RICE, 1979 The kinetics of conjugative plasmid transmission: fit of a simple mass action model. Plasmid **2:** 247–260.

LIM, J. K., 1979 Site-specific instability in *Drosophila melanogaster*: the origin of the mutation and cytogenetic evidence for site specificity. Genetics **93:** 681–701.

LIM, J. K., M. J. SIMMONS, J. D. RAYMOND, N. M. COX, R. F. DOLL and T. P. CULBERT, 1983 Homologue destabilization by a putative transposable element in *Drosophila melanogaster*. Proc. Natl. Acad. Sci. USA **80:** 6624–6627.

LINDSLEY, D. L. and L. SANDLER, 1977 The genetic analysis of meiosis in female *Drosophila melanogaster*. Philos. Trans. R. Soc. Lond. [Biol] **277:** 295–312.

MAYNARD SMITH, J., 1982 *Evolution and the Theory of Games*. Cambridge University Press, Cambridge.

MCCLINTOCK, B., 1956 Controlling elements and the gene. Cold Spring Harbor Symp. Quant. Biol. **21:** 197–216.

MONTGOMERY, E. A. and C. H. LANGLEY, 1983 Transposable elements in Mendelian populations. II. Distribution of three *copia*-like elements in a natural population. Genetics **104:** 473–483.

MORTON, N. E., 1955 Sequential tests for the detection of linkage. Am. J. Hum. Genet. **7:** 277–318

MUKAI, T., 1969 The genetic structure of natural populations of *Drosophila melanogaster*. VII. Synergistic interaction of mutant polygenes controlling viability. Genetics **61:** 749–761.

MULLER, H. J. and T. S. PAINTER, 1932 The differentiation of the sex chromosomes of *Drosophila* into genetically active and inert regions. Z. Indukt. Abstammungs.-Vererbungsl. **62:** 316–365.

OHTA, T., 1981 Population genetics of selfish DNA. Nature **292:** 648–649.

OHTA, T., 1983 Theoretical study on the accumulation of selfish DNA. Genet. Res. **41:** 1–16.

OHTA, T., 1984 Population genetics of transposable elements. IMA J. Math. Appl. Med. Biol. **1:** 17–29.

OHTA, T., 1985 A model of duplicative transposition and gene conversion. Genetics **110:** 513–524.

OHTA, T. and M. KIMURA, 1981 Some calculations on the amount of selfish DNA. Proc. Natl. Acad. Sci. USA **78:** 1129–1132.

ROBERTSON, D. S., 1983 A possible dose-dependent inactivation of mutator (*Mu*) in maize. Mol. Gen. Genet. **191:** 86–90.

ROBINSON, M. K., P. M. BENNETT and M. J. RICHMOND, 1977 Inhibition of TnA translocation by TnA. J. Bacteriol. **129:** 407–414.

ROEDER, G. S. and G. R. FINK, 1983 Transposable elements in yeast. pp. 299–328. In: *Mobile Genetic Elements,* Edited by J. A. SHAPIRO. Academic Press, New York.

ROTH, G. E. and K. B. MORITZ, 1981 Restriction enzyme analysis of the germ line limited DNA of *Ascaris suum.* Chromosoma **83:** 169–190.

RUBIN, G. M., 1983 Dispersed repetitive DNA in *Drosophila.* pp. 329–361. In: *Mobile Genetic Elements,* Edited by J. A. SHAPIRO. Academic Press, New York.

RUBIN, G. M., M. G. KIDWELL and P. M. BINGHAM, 1982 The molecular basis of *P-M* hybrid dysgenesis: the nature of induced mutations. Cell **29:** 987–994.

SCHALET, A. and G. LEFEVRE, 1976 The proximal region of the X chromosome. pp. 847–902. In: *The Genetics and Biology of Drosophila.* Vol. 1, Edited by M. ASHBURNER and E. NOVITSKI. Academic Press, London.

SCHWARTZ, D., 1984 Analysis of the *Ac* element dosage effect in maize. Mol. Gen. Genet. **196:** 81–86.

SHAPIRO, J. A., 1983 *Mobile Genetic Elements.* Academic Press, New York.

SHERRATT, D. J., A. ARTHUR, R. BISHOP, P. DYSON, P. KITTS and L. SYMINGTON, 1983 Genetic transposition in bacteria. pp. 59–74. In: *Genetic Rearrangement,* Edited by K. F. CHATER, C. A. CULLIS, D. A. HOPWOOD, A. W. B. JOHNSTON and H. W. WOOLHOUSE. Croom Helm, London.

SIMMONS, M. J. and J. F. CROW, 1977 Mutations affecting fitness in *Drosophila* populations. Annu. Rev. Genet. **11:** 49–78.

SIMONS, R. W. and N. KLECKNER, 1983 Translational control of IS10 transposition. Cell **34:** 683–691.

SLATKIN, M., 1985 Genetic differentiation of transposable elements under mutation and unbiased gene conversion. Genetics **110:** 145–158.

SYVANEN, M., 1984 The evolutionary implication of mobile genetic elements. Annu. Rev. Genet. **18:** 271–293.

WHITE, M. J. D., 1973 *Animal Cytology and Evolution,* Ed. 3. Cambridge University Press, Cambridge.

YOKOYAMA, R. W. and M.-C YAO, 1982 Elimination of DNA sequences during macronuclear differentiation of *Tetrahymena thermophila,* as detected by in situ hybridization. Chromosoma **85:** 11–22.

Communicating editor: B. S. WEIR

APPENDIX 1. GENERAL CONSIDERATIONS

Consider all the individuals who carry a mutant element in the heterozygous state. We define *cis* sites in such individuals as sites located in chromosomes that entered the zygote in the same gamete as the mutant element and define *trans* sites as those located in the complementary set of chromosomes. Let x_{ic} and x_{it} be the frequencies of nonmutant elements at the ith occupable site, before selection, in *cis* and in *trans*, respectively. Since the mutant is assumed to be rare, x_{it} can be approximated by $\hat{x}$, the equilibrium frequency of the element at each site in the initial population before the introduction of the mutation. Because of the effect of the mutation on transposition, $x_{ic} \neq x_{it}$, and we can write $x_{ic} = \hat{x} + \delta x_i$. The net frequency of elements at the ith site among mutant individuals is thus $x_{im} = \frac{1}{2}(x_{ic} + x_{it}) = \hat{x} + \frac{1}{2}\delta x_i$. The deviation of the mean number of copies of the element per individual among mutant individuals, from the mean $\bar{n}$ for the nonmutant population, is $\delta\bar{n} = \sum_i \delta x_i$. (At first sight, it might seem that the fact that mutant elements are necessarily present in a genome with at least one element would cause them to be associated with a higher mean copy number than average. But if $\bar{n}$ is sufficiently large, the frequency of individuals lacking any elements is so small that this effect is negligible; *e.g.*, with $\bar{n} = 10$, this frequency is $\exp -10 \simeq 4.5 \times 10^{-5}$.)

After selection in a given generation, these frequencies are modified to

$$x_{ic}^* = x_{ic} + \frac{x_{ic}}{\bar{w}_m}\frac{\partial \bar{w}_m}{\partial x_{it}} \tag{A.1a}$$

$$x_{it}^* = x_{it} + \frac{x_{it}}{\bar{w}_m}\frac{\partial \bar{w}_m}{\partial x_{it}}, \tag{A.1b}$$

where $\bar{w}_m$ is the mean fitness of the mutant individuals.

From the definition of x_{im}, we obtain

$$\bar{w}_m \simeq \bar{w} + \tfrac{1}{2}\sum_j \delta x_j \frac{\partial \bar{w}}{\partial \hat{x}} \tag{A.2a}$$

$$\frac{\partial \bar{w}_m}{\partial x_{ic}} = \frac{\partial \bar{w}_m}{\partial x_{it}} = \frac{1}{2}\frac{\partial \bar{w}_m}{\partial x_{im}} \simeq \frac{1}{2}\frac{\partial \bar{w}}{\partial \hat{x}} + \frac{1}{4}\sum_j \delta x_j \frac{\partial^2 \bar{w}}{\partial \hat{x}^2}, \tag{A.2b}$$

where $\bar{w}$ is the mean fitness of the nonmutant population. Noting that $\partial\bar{w}/\partial\bar{n} = \frac{1}{2}\partial\bar{w}/\partial\hat{x}$, these can be rewritten more conveniently as

$$\bar{w}_m \simeq \bar{w} + \delta\bar{n}\frac{\partial \bar{w}}{\partial \bar{n}} \tag{A.3a}$$

$$\frac{\partial \bar{w}}{\partial x_{ic}} = \frac{\partial \bar{w}}{\partial x_{it}} \simeq \frac{\partial \bar{w}}{\partial \bar{n}} + \delta\bar{n}\frac{\partial^2 \bar{w}}{\partial \bar{n}^2}. \tag{A.3b}$$

Substituting into (A.1), and neglecting second-order terms in δx_i and $\partial\bar{w}/\partial\bar{n}$, we obtain

$$x_{ic}^* \simeq (\hat{x} + \delta x_i)\left(1 + \frac{\partial ln\ \bar{w}}{\partial \bar{n}}\right) + \frac{\hat{x}\delta\bar{n}}{\bar{w}}\left(\frac{\partial^2 \bar{w}}{\partial \bar{n}^2}\right) \tag{A.4a}$$

$$x_{it}^* \simeq \hat{x}\left(1 + \frac{\partial ln\ \bar{w}}{\partial \bar{n}}\right) + \frac{\hat{x}\delta\bar{n}}{\bar{w}}\left(\frac{\partial^2 \bar{w}}{\partial \bar{n}^2}\right). \tag{A.4b}$$

For an equilibrium population, we have $\partial ln\ \bar{w}/\partial\bar{n} = v - u$ [see (1) of the text]. Furthermore, if copy number before selection is approximately normally distributed with variance $V_n \simeq \bar{n}$ (CHARLESWORTH and CHARLESWORTH 1983) and the variance after selection is V_n^*, we can write

$$k = \frac{V_n^* - V_n}{V_n} \simeq \frac{V_n^* - V_n}{\bar{n}} = \frac{\bar{n}}{\bar{w}}\frac{\partial^2 \bar{w}}{\partial \bar{n}^2}. \tag{A.5}$$

Thus, we obtain the relations

$$x_{ic}^{*} \simeq (\hat{x} + \delta x_i)(1 + v - u) + \frac{k\delta\bar{n}}{2m} \tag{A.6a}$$

$$x_{il}^{*} \simeq \hat{x}(1 + v - u) + \frac{k\delta\bar{n}}{2m}. \tag{A.6b}$$

Further development of these relations, to obtain the values of the x_{ic} and, hence, $\delta\bar{n}$ in the next generation, varies according to the specific model of transposition regulation in question.

APPENDIX 2. TRANSPOSITION IMMUNITY

We assume that the mutant element prevents the insertion of elements into a region ρ map units long, centered on its own site of insertion. The average rate of transposition of all elements in a genome carrying the mutant is changed by δu_a; the rate of transposition of the mutant itself is changed by $\delta u_m (|\delta u_m| \leq |\delta u_a|$; see text). Let r_i be the frequency of recombination between the mutant element and the ith occupable site. For $r_i < \frac{1}{2}\rho$, the model implies that x_{ic}^{*} in (A.6a, b) is not changed by transposition. For elements with $r_i > \frac{1}{2}\rho$, the value of x_{ic}^{*} is modified as follows. One-half of all new transpositions will be inserted in coupling with the mutant; the total number of these is thus $\frac{1}{2}(\bar{n} + \delta\bar{n})(u + \delta u_a)$, and their frequency of insertion into the ith site is $1/m$ if m is large. Neglecting second-order terms in $\delta\bar{n}$ and δu_a, the net transpositional contribution to x_{ic}, the value of x_{ic} in the next generation, is thus

$$\frac{\bar{n}(u + \delta u_a)}{2m} + \frac{u\delta\bar{n}}{2m} = \hat{x}(u + \delta u_d) + \frac{u\delta\bar{n}}{2m} \qquad (r_i > \tfrac{1}{2}\rho). \tag{A.7}$$

Taking recombination between coupling and repulsion sites into account, the net contribution of excision events to x'_{ic} is similarly

$$-v(\hat{x} + \delta x_i[1 - r_i]). \tag{A.8}$$

(No restriction is placed on r_i in this equation.) The final equations for x'_{ic}, allowing for recombination, are

$$x'_{ic} \simeq \hat{x}(1 - u) + \frac{k\delta\bar{n}}{2m} + \delta x_i(1 - r_i)(1 - u) \qquad (r_i \leq \tfrac{1}{2}\rho) \tag{A.9a}$$

$$x'_{ic} \simeq \hat{x}(1 + \delta u_a) + \frac{(u + k)\delta\bar{n}}{2m} + \hat{x}\delta u_a + \delta x_i(1 - r_i)(1 - u) \qquad (r_i > \tfrac{1}{2}\rho). \tag{A.9b}$$

Subtracting $\hat{x}$ from both sides, we obtain

$$\delta x'_i \simeq \delta x_i(1 - r_i)(1 - u) + \frac{k\delta\bar{n}}{2m} - \hat{x}u \qquad (r_i \leq \tfrac{1}{2}\rho) \tag{A.10a}$$

$$\delta x'_i \simeq \delta x_i(1 - r_i)(1 - u) + \frac{(u + k)\delta\bar{n}}{2m} + \hat{x}\delta u_a \qquad (r_i > \tfrac{1}{2}\rho). \tag{A.10b}$$

The asymptotic values of the $\delta x_i (\delta\hat{x}_i)$ are given by the relations

$$\delta\hat{x}_i(r_i[1 - u] + u) = \frac{k\delta\bar{n}}{2m} - \frac{\bar{n}u}{2m} \qquad (r_i \leq \tfrac{1}{2}\rho) \tag{A.11a}$$

$$\delta\hat{x}_i(r_i[1 - u] + u) = \frac{(u + k)\delta\bar{n}}{2m} + \frac{\bar{n}\delta u_a}{2m} \qquad (r_i > \tfrac{1}{2}\rho). \tag{A.11b}$$

We note that k is generally of order u or less; for a selection function of the form $w_n =$

exp $-tn^2$, $k \simeq u - v$ for an equilibrium population, and with truncation selection $k \simeq \bar{n}(u - v)^2$ (*cf.* BULMER 1980, p. 153). Since $\delta\bar{n}$ is a small quantity itself, we can neglect the terms $k\delta\bar{n}$ and $(u + k)\delta\bar{n}$ on the right-hand side of (A.11a, b). Rearranging, approximating $r_i(1 - u) + u$ by $r_i + u$, and summing over i, we obtain

$$-\delta\bar{n} \simeq \frac{\bar{n}}{2m}(u\Sigma_1 - \delta u_a \Sigma_2) \tag{A.12}$$

where Σ_1 and Σ_2 are defined by (4) of the text.

Having obtained an expression for the difference in mean copy number between mutant and wild-type genomes, it is straightforward to determine the asymptotic change in frequency per generation of the mutant. Let the frequency of mutant elements at the ith occupable site be y_i. The response of population mean copy number to a small change dy_i in y_i is

$$d\bar{n} = 2dy_i(1 + \delta\bar{n}) \tag{A.13}$$

where (asymptotically) $\delta\bar{n}$ is given by (A.12). (This follows from the fact that $2dy_i$ is the change in mean copy number arising purely from an increase in element frequency, and $2dy_i\delta\bar{n}$ is the change arising from the difference in mean copy number associated with the mutant.) Hence,

$$\frac{1}{2}\frac{\partial ln\ \bar{w}}{\partial y_i} = \frac{1}{2}\frac{\partial ln\ \bar{w}}{\partial \bar{n}}\frac{\partial \bar{n}}{\partial y_i} = (1 + \delta\bar{n})\frac{\partial ln\ \bar{w}}{\partial \bar{n}}. \tag{A.14}$$

Summing over all sites, and using equation (18a) of CHARLESWORTH and CHARLESWORTH (1983) for large m, we obtain the change in net frequency of the mutant element as

$$\sum_i \Delta y_i \simeq \left(\sum_i y_i\right)\left\{(1 + \delta\bar{n})\frac{\partial ln\ \bar{w}}{\partial \bar{n}} + u + \delta u_m - v\right\}. \tag{A.15}$$

Using (1) for the initial equilibrium population, this gives the selection coefficient s on the mutant element as

$$s = \frac{\left(\sum_i \Delta y_i\right)}{\left(\sum_i y_i\right)} \simeq -\delta\bar{n}(u - v) + \delta u_m, \tag{A.16}$$

where $\delta\bar{n}$ is given by (A.12).

APPENDIX 3. TRANSPOSITION REPRESSION MODEL

In this model, we assume that the mutant element changes the rate of transposition of all elements (including itself) in the same genome by δu_a. The argument leading to (A.10b) above is still valid, except that there is no longer any restriction $r_i > \frac{1}{2}\rho$. Instead of (A.12), we obtain the asymptotic relation

$$-\delta\bar{n} \simeq \left\{\sum_i \frac{1}{(r_i + u)}\right\}\frac{\bar{n}\delta u_a}{2m}, \tag{A.17}$$

which can be rewritten as

$$-\delta\bar{n} \simeq \frac{\bar{n}\delta u_a}{2H}, \tag{A.18}$$

where $H = m/\Sigma\dfrac{1}{(r_i + u)}$ is the harmonic mean of $r_i + u$ over all pairs of loci.

Since $\delta u_m = \delta u_a$ in this case, the selection coefficient on the mutant is now

$$s \simeq -\delta u_a \left\{ \frac{\bar{n}(u - v)}{2H} - 1 \right\}. \tag{A.19}$$

APPENDIX 4. CONDITIONAL TRANSPOSITION REPRESSION MODEL

Transposition is assumed here to be prevented by the mutant element in genomes where the copy number n exceeds a threshold value n_c. Assuming that copy number among mutant individuals is asymptotically normally distributed, with mean and variance $\bar{n} + \delta\bar{n}$, we can write $\phi(z)$ for the corresponding standardized normal variate, where $z = (n - [\bar{n} + \delta\bar{n}])(\bar{n} + \delta\bar{n})^{-\frac{1}{2}}$. Let $\phi_c = \phi(z_c)$, where z_c corresponds to $n = n_c$, and let Φ_c be the corresponding cumulative probability.

The mean rate of transposition of mutant elements is then

$$u_m \simeq u \int_{-\infty}^{z_c} \phi(z)dz = u\Phi_c \tag{A.20}$$

so that

$$\delta u_m \simeq -u(1 - \Phi_c). \tag{A.21}$$

The mean number of transposition events involving all elements in mutant genomes is given approximately by

$$u\{(\bar{n} + \delta\bar{n}) + (\bar{n} + \delta\bar{n})^{\frac{1}{2}} \int_{-\infty}^{z_c} z\phi(z)dz\} = u(\bar{n} + \delta\bar{n})\{\Phi_c - (\bar{n} + \delta\bar{n})^{-\frac{1}{2}}\phi_c\}, \tag{A.22}$$

and so the change in mean transposition rate for all elements is

$$\delta u_a \simeq -u(1 - \Phi_c + \bar{n}^{\frac{1}{2}}\phi_c), \tag{A.23}$$

where $\bar{n} + \delta\bar{n}$ in the second term of the right-hand side of (A.22) has been approximated by $\bar{n}$.

By the same method as before, the selection coefficient on a mutant element is

$$s = u(1 - \Phi_c)\left\{ \frac{\bar{n}(u - v)F}{2H} - 1 \right\}, \tag{A.24}$$

where

$$F = (1 - \Phi_c + \bar{n}^{\frac{1}{2}}\phi_c)/(1 - \Phi_c).$$

APPENDIX 5. DOMINANT LETHALITY MODEL

We here use the transposition repression model with the additional assumption that a proportion π of transpositions result in a dominant lethal or sterile mutation that causes the genetic death of the progeny individuals who carry them. Using the formulas developed in the text, we can replace u in (A.19) by the effective transposition rate $\tilde{u} = u(1 - 2\pi)$, δu_a by $\delta\tilde{u}_a = \delta u(1 - 2\pi)$, and H by the harmonic mean of $r_i + \tilde{u}$ ($\tilde{H}$). The selection coefficient obtained from this formula, however, omits the fact that mutant genomes have a lower probability of carrying a newly induced dominant lethal mutation than do normal individuals. This probability for mutant individuals is $\frac{1}{2}(\bar{n} + \delta\bar{n})\pi(u + \delta u_a) + \frac{1}{2}\bar{n}\pi u$, compared with a probability of $\bar{n}\pi u$ for a normal individual. Ignoring second-order terms, the mutant individuals thus obtain a selective advantage of $-\frac{1}{2}\bar{n}\pi\delta u_a$.

The net selection coefficient on a mutant element is thus

$$s \simeq -\delta\tilde{u}_a \left\{ \frac{\bar{n}(\tilde{u} - v)}{2\tilde{H}} + \frac{\bar{n}\pi}{2(1 - 2\pi)} - 1 \right\}. \tag{A.25}$$

APPENDIX 6. EFFECTS OF HAPLOIDY

A life cycle with an extended haploid vegetative phase will be assumed, with a transient diploid zygote followed by meiosis. The basic selection equations (A.1) to (A.6) are unchanged, except that we need consider only the x_{ic}; there are now no effects of altered selection and transposition rates on the x_{it} due to association with mutant elements in repulsion. Because of haploidy, we have $\hat{x} = \bar{n}/m$, where $\bar{n}$ is now the mean for haploid genomes. Applying similar reasoning to that used to obtain (A.6a), we find

$$x_{ic}^{*} = (\hat{x} + \delta x_i)(1 + v - u) + \frac{k\delta\bar{n}}{m}. \tag{A.26}$$

In order to model the consequences of altered transposition rates, we assume that transposition and excision take place during the haploid phase, so that all new insertions occur in coupling with the mutant element. For the transposition immunity model, we now find that

$$\delta x_i' \simeq \delta x_i(1 - r_i)(1 - u) + \frac{k\delta\bar{n}}{m} - \hat{x}u \qquad (r_i \leqslant 1/2\rho) \tag{A.27a}$$

$$\delta x_i' \simeq \delta x_i(1 - r_i)(1 - u) + \frac{(u + k)\delta\bar{n}}{m} + (1 - r_i)\hat{x}\delta u_a \ (r_i > 1/2\rho). \tag{A.27b}$$

The term $(1 - r_i)\hat{x}\delta u_a$ replaces $x\delta u_a$ in (A.10b), because only $(1 - r_i)$ of the new insertions remain in coupling with the mutant after meiosis has taken place. Equation (A.12) is therefore replaced by

$$\delta\bar{n} \simeq -\frac{\bar{n}}{m}(u\Sigma_1 - \delta u_a\Sigma_2) \tag{A.28}$$

where $\Sigma_1 = \sum_{r_i=0}^{1/2\rho} \frac{1}{(r_i + u)}$, as before, but now $\Sigma_2 = \sum_{r_i>1/2\rho}^{1/2} \frac{(1 - r_i)}{(r_i + u)}$.

The selection coefficient on the mutant element is calculated in the same way as before [(A.16)].

A similar procedure can be followed for the case of the transposition repression models. It is easily shown that H in (A.18) is replaced by

$$H^{*} = m/\sum_i \frac{(1 - r_i)}{(r_i + u)} \tag{A.29}$$

and that

$$s \simeq -\delta u_a \left\{\frac{\bar{n}(u - v)}{H^{*}} - 1\right\}. \tag{A.30}$$

The dominant lethality model is replaced by a model that assumes that a fraction π of the transpositions occurring in the vegetative phase are lethal. Since all lethal transpositions occurring in a mutant genome result in elimination of the mutant element, the direct selective advantage of the mutant is now $-\bar{n}\pi\delta u_a$; therefore, (A.25) is replaced by

$$s \simeq -\delta\tilde{u}_a \left\{\frac{\bar{n}(u - v)}{\tilde{H}^{*}} + \frac{\bar{n}\pi}{(1 - 2\pi)} - 1\right\}, \tag{A.31}$$

where $\tilde{H}^{*}$ is the harmonic mean of $(r_i + \tilde{u})/(1 - r_i)$.